Campus de l'Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7 Canada

Telephone (514) 343-7575
Télécopieur (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

CRT-2002-09

Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

GENERALIZED NETWORK DESIGN PROBLEMS

by

Corinne Feremans
Martine Labbé
Gilbert Laporte

April 2002

PUBLICATION
CRT-2002-09

GENERALIZED NETWORK DESIGN PROBLEMS*

by

Corinne Feremans[1,2]
Martine Labbé[1]
Gilbert Laporte[3]

* This research was supported in part by the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture, by the Région Bruxelles-Capitale, by the Direction générale des Relations Extérieures de la Région wallone de Belgique and by the Natural Sciences and Engineering Research Council of Canada (NSERC).

1 Institut de Statistique et de Recherche Opérationnelle, Service d'Optimisation, C.P. 210/01, Université Libre de Bruxelles, boulevard du Triomphe, B-1050 Bruxelles, Belgium (mlabbe@smg.ulb.ac.be)

2 Universiteit Maastricht, Faculty of Economics and Business Administration Department, Quantitative Economics, P.O. Box 616, 6200 MD Maastricht, The Netherlands (C.Feremans@KE.unimaas.nl)

3 Canada Research Chair in Distribution Management, École des Hautes Études Commerciales, 3000 chemin de la Côte-Sainte-Catherine, Montréal H3T 2A7 Canada and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-Ville, Montréal H3C 3J7 Canada (gilbert@crt.umontreal.ca)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2002

103-03056/

Abstract

Network design problems consists of identifying an optimal subgraph of a graph, subject to side constraints. In generalized network design problems, the vertex set is partitioned into clusters and the feasibility conditions are expressed in terms of the clusters. Several applications of generalized network design problems arise in the fields of telecommunications, transportation and biology. The aim of this review article is to formally define generalized network design problems, to study their properties and to provide some applications.

Key words: generalized network design problems, complexity.

Résumé

Les problèmes de conception de réseaux consiste à identifier un sous-graphe optimal d'un graphe satisfaisant à des contraintes subsidiaires. Dans les problèmes de conception de réseaux généralisés, les conditions d'admissibilité s'expriment en termes de sous-ensembles de sommets. On retrouve plusieurs applications des problèmes de conception de réseaux généralisés en télécommunications, en transport et en biologie. Le but de cet article-synthèse est de définir les problèmes de conception de réseaux généralisés de façon formelle, d'étudier leurs propriétés et de fournir quelques applications.

Mots-clefs : problème de conception de réseaux généralisés, complexité.

1 Introduction

Several classical combinatorial optimization problems can be cast as *Network Design Problems (NDP)*. Broadly speaking, an NDP consists of identifying an optimal subgraph F of an undirected graph G subject to feasibility conditions. Well known NDPs are the Minimum Spanning Tree Problem (MSTP), the Traveling Salesman Problem (TSP) and the Shortest Path Problem (SPP). We are interested here in *Generalized NDPs*, i.e., in problems where the vertex set of G is partitioned into *clusters* and the feasibility conditions are expressed in terms of the clusters. For example, one may wish to determine a minimum length tree spanning all the clusters, a Hamiltonian cycle through all the clusters, etc.

Generalized NDPs are important combinatorial optimization problems in their own right, not all of which have received the same degree of attention by operational researchers. In order to solve them, it is useful to understand their structure and to exploit the relationships that link them. These problems also underlie several important applications areas, namely in the fields of telecommunications, transportation and biology.

Our aim is to formally define generalized NDPs, to study their properties and to provide examples of their applications. We will first define an unified notational framework for these problems. This will be followed by complexity results and by the study of seven generalized NDPs.

2 Definitions and notations

An undirected *graph* $G = (V, E)$ consists of a finite non-empty vertex set $V = \{1, \ldots, n\}$ and an edge set $E \subseteq \{\{i, j\} : i, j \in V\}$. Costs c_i and c_{ij} are assigned to vertices and edges respectively. Unless otherwise specified, $c_i = 0$ for $i \in V$ and $c_{ij} \geq 0$ for $\{i, j\} \in E$. We denote by $E(S) = \{\{i, j\} \in E : i, j \in S\}$, the subset of edges having their two end vertices in $S \subseteq V$. A subgraph F of G is denoted

by $F = (V_F, E_F), V_F \subseteq V, E_F \subseteq E(V_F)$, and its cost $c(F)$ is the sum of its vertex and edge costs. It is convenient to define an NDP as a problem P associated with a subset of *terminal vertices* $T \subseteq V$. A feasible solution to P is a subgraph $F = (V_F, E_F)$, where $T \subseteq V_F$, satisfying some side constraints. If $T = V$, then the NDP is *spanning*; if $T \subset V$, it is *non-spanning*. Let $G(T) = (T, E(T))$ and denote by $\mathcal{F}_P(T)$ the subset of feasible solutions to the spanning problem P defined on the graph $G(T)$. Let $S \subseteq V$ be such that $S \cap T = \emptyset$, and denote by $\mathcal{F}_P(T, S)$ the set of feasible solutions of the non-spanning problem P on graph $G(S \cup T)$ that spans T, and possibly some vertices from S.

In this framework, feasible NDP solutions correspond to a subset of edges satisfying some constraints. Natural spanning NDPs are the following.

1. The *Minimum Spanning Tree Problem* (MSTP) (see e.g., Magnanti and Wolsey [45]). The MSTP is to determine a minimum cost tree on G that includes all the vertices of V. This problem is polynomially solvable.

2. The *Traveling Salesman Problem* (TSP) (see e.g., Lawler, Lenstra, Rinnooy Kan and Shmoys [42]). The TSP consists of finding a minimum cost cycle that passes through each vertex exactly once. This problem is $\mathcal{NP}$-hard.

3. The *Minimum Perfect Matching Problem* (MPMP) (see e.g., Cook, Cunningham, Pulleyblank and Schrijver [8]). A matching $M \subseteq E$ is a subset of edges such that each vertex of M is adjacent to at most one edge of M. A perfect matching is a matching that contains all the vertices of G. The problem consists of finding a perfect matching of minimum cost. This problem is polynomial.

4. The *Minimum 2-Edge-Connected Spanning Network* (M2ECN) (see e.g., Grötschel, Monma and Stoer [26] and Mahjoub [46]. The M2ECN consists of finding a subgraph with minimal total cost for which there exists two edge-disjoint paths between every pair of vertices.

5. The *Minimum Clique Problem* (MCP). The MCP consists of determining a minimum total cost clique spanning all the vertices. This problem is trivial since the whole graph corresponds to an optimal solution.

We also consider the following two non-spanning NDPs.

1. The *Steiner Tree Problem* (STP) (see Winter [61] for an overview). The STP is to determine a tree on G that spans a set T of terminal vertices at minimum cost. A *Steiner tree* may contain vertices other than those of T. These vertices are called the *Steiner vertices*. This problem is $\mathcal{NP}$-hard.

2. The *Shortest Path Problem* (SPP) (see e.g., Ahuja, Magnanti and Orlin [1]). Given an origin o and a destination d, $o, d \in V$, the SPP consists of determining a path of minimum cost from o to d. This problem is polynomially solvable. It can be seen as a particular case of the STP where $T = \{o, d\}$.

In generalized NDPs, V is partitioned into clusters V_k, $k \in K$. We now formally define spanning and non-spanning generalized NDPs.

Definition 1 ("Exactly" generalization of spanning problem). *Let $G = (V, E)$ be a graph partitioned into clusters V_k, $k \in K$. The* "exactly" generalization *of a spanning NDP P on G consists of identifying a subgraph $F = (V_F, E_F)$ of G yielding*

$$\min\{c(F) : |V_F \cap V_k| = 1, F \in \mathcal{F}_P(\bigcup_{k \in K}(V_F \cap V_k))\}.$$

In other words, F must contain exactly one vertex per cluster. Two different generalizations are considered for non-spanning NDPs.

Definition 2 ("Exactly" generalizations of non-spanning problem). *Let $G = (V, E)$ be a graph partitioned into clusters V_k, $k \in K$, and let $\{K_T, K_S\}$ be a partition of K. The* "exactly" T-generalization *of a non-spanning problem NDP P on G consists of identifying a subgraph $F = (V_F, E_F)$ of G yielding*

$$\min\{c(F) : |V_F \cap V_k| = 1, k \in K_T, F \in \mathcal{F}_P(\bigcup_{k \in K_T}(V_F \cap V_k), \bigcup_{k \in K_S} V_k)\}.$$

The "exactly" S-generalization *of a non-spanning problem NDP P on G consists of identifying a subgraph $F = (V_F, E_F)$ of G yielding*

$$\min\{c(F) : |V_F \cap V_k| = 1, k \in K_S, F \in \mathcal{F}_P(\bigcup_{k \in K_T} V_k, \bigcup_{k \in K_S}(V_F \cap V_k))\}.$$

In other words, in the *"exactly" T-generalization*, F must contain exactly one vertex per cluster V_k with $k \in K_T$, and possibly other vertices in $\bigcup_{k \in K_S} V_k$. In the *"exactly" S-generalization*, F must contain exactly one vertex per cluster V_k with $k \in K_S$, and all vertices of $\bigcup_{k \in K_T} V_k$.

We can replace $|V_F \cap V_k| = 1$ in the above definitions by $|V_F \cap V_k| \geq 1$ or $|V_F \cap V_k| \leq 1$, leading to the "at least" version or "at most " version of the generalization. The "exactly", "at least" and "at most" versions of a generalized NDP P are denoted by E-P, L-P and M-P, respectively. In the "at most" and in the "exactly" versions, intra-cluster edges are neglected. In this case, we call the graph G, *$|K|$-partite complete*. In the "at least" version the intra-cluster edges are taken into account.

3 Complexity results

We provide in Tables 1 and 2 the complexity of the generalized versions in their three respective forms ("exactly", "at least" and "at most") for the seven NDPs considered. Some of these combinations lead to trivial problems. Obviously, if a classical NDP is $\mathcal{NP}$-hard, its generalization is also $\mathcal{NP}$-hard. The indication "$\emptyset$ is opt" means that the empty set is feasible and is optimal for the corresponding problem. References about complexity results for the classical version of the seven problems considered can be found in Garey and Johnson [20].

As can be seen from Table 2, two cases of the generalized SPP are $\mathcal{NP}$-hard by reduction from the *Hamiltonian Path Problem* (see Garey and Johnson [20]). Li, Tsao and Ulular [43] show that the "at most" S-generalization is polynomial if the *shrunk graph* is *series-parallel* but provide no complexity result for the general case. A shrunk graph $G_S = (V_S, E_S)$ derived from a graph G partitioned into clusters is defined as follows: V_S contains one vertex for each cluster of G, and there exists an edge in E_S whenever an edge between the two corresponding clusters exists in G. An undirected graph is series-parallel if it is not *contractible* to K_4, the complete graph on four vertices. A graph G is contractible to another graph H if H can be obtained from G by deleting and contracting edges. Contracting an edge means that its two end vertices are shrunk and the edge is deleted.

We now provide a short literature review and applications for each of the seven generalized NDPs considered.

Table 1: Complexity of classical and generalized spanning NDPs

Problem	MSTP	TSP	MPMP	M2ECN	MCP
Classical	Polynomial	$\mathcal{NP}$-hard	Polynomial	$\mathcal{NP}$-hard	Trivial, polynomial
Exactly	$\mathcal{NP}$-hard [47]	$\mathcal{NP}$-hard	Polynomial	$\mathcal{NP}$-hard	$\mathcal{NP}$-hard (with vertex cost)[35]
At least	$\mathcal{NP}$-hard [31]	$\mathcal{NP}$-hard	Polynomial	$\mathcal{NP}$-hard	Equivalent to exactly
At most	$\emptyset$ is opt	$\emptyset$ is opt	$\emptyset$ is opt	$\emptyset$ is opt	$\emptyset$ is opt

Table 2: Complexity of classical and generalized non-spanning NDPs

Problem	STP	SPP
Classical	$\mathcal{NP}$-hard	Polynomial
Exactly T-generalization	$\mathcal{NP}$-hard	Polynomial
Exactly S-generalization	$\mathcal{NP}$-hard	$\mathcal{NP}$-hard
At least T-generalization	$\mathcal{NP}$-hard	Polynomial
At least S-generalization	$\mathcal{NP}$-hard	$\mathcal{NP}$-hard
At most T-generalization	$\emptyset$ is opt	$\emptyset$ is opt
At most S-generalization	$\mathcal{NP}$-hard	Polynomial if shrunk graph is series-parallel [43]

4 The generalized minimum spanning tree problem

The *Generalized Minimum Spanning Tree Problem* (E-GMSTP) is the problem of finding a minimum cost tree including *exactly* one vertex from each vertex set from the partition (see Figure 1a for a feasible E-GMSTP solution). This problem was introduced by Myung, Lee and Tcha [47]. Several formulations are available for the E-GMSTP (see Feremans, Labbé and Laporte [17]).

The *Generalized Minimum Spanning Tree Problem* in its "at least" version (L-GMSTP) is the problem of finding a minimum cost tree including *at least* one vertex from each vertex set from the partition (see Figure 1b for a feasible solution of L-GMSTP). This problem was introduced by Ihler, Reich and Widmayer [31] as a particular case of the *Generalized Steiner Tree Problem* (see Section 9) under the name *"Class Tree Problem"*. Dror, Haouari and Chaouachi [11] show that if the family of clusters covers V without being pairwise disjoint, then the L-GMSTP defined on this family can be transformed into the original L-GMSTP on a graph G' obtained by substituting each vertex $v \in \bigcap_{\ell \in L} V_\ell$, $L \subseteq K$ by $|L|$ copies $v_\ell \in V_\ell$, $\ell \in L$, and adding edges of weight zero between each pair of these new vertices (clique of weight zero between v_ℓ for $\ell \in L$). This can be done as long as there is no fixed cost on the vertices, and this transformation does not hold for the "exactly" version of the problem.

Applications modeled by the E-GMSTP are encountered in telecommunications, where metropolitan and regional networks must be interconnected by a tree containing a gateway from each network. For this internetworking, a vertex has to be chosen in each local network as a hub and the hub vertices must be connected via transmission links such as optical fiber (see Myung, Lee and Tcha [47]).

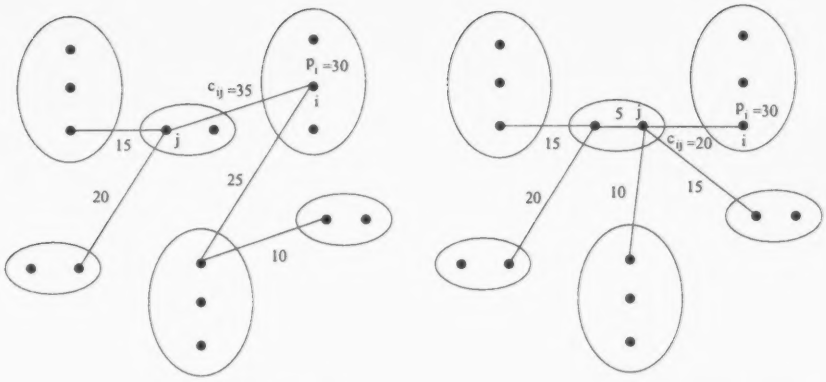

Figure 1a: E–GMSTP

Figure 1b: L–GMSTP

Figure 1: Feasible GMSTP solutions

The L-GMSTP has been used to model and solve an important irrigation network design problem arising in desert environments, where a set of $|K|$ polygon shaped parcels share a common source of water. Each parcel is represented by a cluster made up of the polygon vertices. Another cluster corresponds to the water source vertex. The problem consists of designing a minimal length irrigation network connecting at least one vertex from each parcel to the water source. This irrigation problem can be modeled as an L-GMSTP as follows. Edges correspond to the boundary lines of the parcel. The aim is to construct a minimal cost tree such that each parcel has at least one irrigation source (see Dror, Haouari and Chaouachi [11]).

Myung, Lee and Tcha [47] show that the E-GMSTP is strongly $\mathcal{NP}$-hard, using a reduction from the *Node Cover Problem* (see Garey and Johnson [20]). These authors also provide four integer linear programming formulations. A branch-and-bound method is developed and tested on instances involving up to 100 vertices. For instances containing between 120 and 200 vertices, the method is stopped before the first branching. The lower bounding procedure is a heuristic method which approximates the linear relaxation associated with the dual of a multicommodity flow formulation for the E-GMSTP. A heuristic algorithm finds a primal feasible solution for the E-GMSTP using the lower bound. The branching strategy performed in this method is described in Noon and Bean [48]. A cluster is first selected and branching is performed on each vertex of this cluster.

In Faigle, Kern, Pop and Still [14], another mixed integer formulation for the E-GMSTP is given. The linear relaxation of this formulation is computed for a set of 12 instances containing up to 120 vertices. This seems to yield an

optimal E-GMSTP solution for all but one instance. The authors also use the subpacking formulation from Myung, Lee and Tcha [47] in which the integrality constraints are kept and the subtour constraints are added dynamically. Three instances containing up to 75 vertices are tested.

A branch-and-cut algorithm for the same problem is described in Feremans [15]. Several families of valid inequalities for the E-GMSTP are introduced and some of these are proved to be facet defining. Computational results show that instances involving up to 200 vertices can be solved to optimality using this method. A comparison with the computational results obtained in Myung, Lee and Tcha [47] shows that the gap between the lower bound and the upper bound obtained before branching is reduced by 10 % to 20 %.

Pop, Kern and Still [51] provide a polynomial approximation algorithm for the E-GMSTP. Its worst-case ratio is bounded by 2ρ if the cluster size is bounded by ρ. This algorithm is derived from the method described in Magnanti and Wolsey [45] for the *Vertex Weighted Steiner Tree Problem* (see Section 9).

Ihler, Reich, Widmayer [31] show that the decision version of the L-GMSTP is $\mathcal{NP}$-complete even if G is a tree. They also prove that no constant worst-case ratio polynomial-time algorithm for the L-GMSTP exists unless $\mathcal{P} = \mathcal{NP}$, even if G is a tree on V with edge lengths 1 and 0. They also develop two polynomial-time heuristics, tested on instances up to 250 vertices. Finally, Dror, Haouari and Chaouachi [11] provide three integer linear programming formulations for the L-GMSTP, two of which are not valid (see Feremans, Labbé and Laporte [16]). The authors also describe five heuristics including a genetic algorithm. These heuristics are tested on 20 instances up to 500 vertices. The genetic algorithm performs better than the other four heuristics. An exact method is described in Feremans [15] and compared to the genetic algorithm in Dror, Haouari and Chaouachi [11]. These results show that the genetic algorithm is time consuming compared to the exact approach of Feremans [15]. Moreover the gap between the upper bound obtained by the genetic algorithm and the optimum value increases as the size of the problem becomes larger.

5 The generalized traveling salesman problem

The *Generalized Traveling Salesman Problem*, denoted by E-GTSP, consists of finding a least cost cycle passing through each cluster *exactly* once. The symmetric E-GTSP was introduced by Henry-Labordere [28], Saskena [56] and Srivastava, Kumar, Garg and Sen [60] who proposed dynamic programming formulations. The first integer linear programming formulation is due to Laporte and Nobert [40] and was later enhanced by Fischetti, Salazar and Toth [18] who in-

troduced a number of facet defining valid inequalities for both the E-GTSP and the L-GTSP. In Fischetti, Salazar and Toth [19], a branch-and-cut algorithm is developed, based on polyhedral results developed in Fischetti, Salazar and Toth [18]. This method is tested on instances whose edge costs satisfy the triangular inequality (for which E-GTSP and L-GTSP are equivalent). Moreover heuristics producing feasible E-GTSP solutions are provided.

Noon [50] has proposed several heuristics for the GTSP. The most sophisticated heuristic published to date is due to Renaud and Boctor [53]. It is a generalization of the heuristic proposed in Renaud, Boctor and Laporte [54] for the classical TSP. Snyder and Daskin [59] have developed a genetic algorithm which is compared to the branch-and-cut algorithm of Fischetti, Salazar and Toth [19] and to the heuristics of Noon [50] and of Renaud and Boctor [53]. This genetic algorithm is slightly slower than other heuristics, but competitive with the CPU times obtained in Fischetti, Salazar and Toth [19] on small instances, and noticeably faster on the larger instances (containing up to 442 vertices).

Approximation algorithms for the GTSP with cost function satisfying the triangle inequality are described in Slavík [58] and in Garg, Konjevod and Ravi [21]. A non-polynomial-time approximation heuristic derived from Christofides heuristic for the TSP [7] is presented in Dror and Haouari [10]; it has a worst-case ratio of 2.

Transformations of the GTSP instances into TSP instances are studied in Dimitrijević and Saric [9], Laporte and Semet [41], Lien, Ma and Wah [44], Noon and Bean [49]. According to Laporte and Semet [41], they do not provide any significant advantage over a direct approach since the TSP resulting from the transformation is highly degenerate.

The GTSP arises in several application contexts, several of which are described in Laporte, Asef-Vaziri and Sriskandarajah [38]. These are encountered in post box location (Labbé and Laporte [36]) and in the design of postal delivery routes (Laporte, Chapleau, Landry, and Mercure [39]). In the first problem the aim is to select a post box location in each zone of a territory in order to achieve a compromise between user convenience and mail collection costs. In the second application, collection routes must be designed through several post boxes at known locations. Asef-Vaziri, Laporte, and Sriskandarajah [3] study the problem of optimally designing a loop-shaped system for material transportation in a factory. The factory is partitioned into $|K|$ rectilinear zones and the loop must be adjacent to at least one side of each zone, which can be formulated as a GTSP. The GTSP can also be used to model a simple case of the stochastic vehicle routing problem with recourse (Dror, Laporte and Louveaux [12]) and some families of arc routing problems (Laporte [37]). In the latter application, a

symmetric arc routing problem is transformed into an equivalent vertex routing problem by replacing edges by vertices. Since the distance from edge e_1 to edge e_2 depends on the traversal direction, each edge is represented by two vertices, only one of which is used in the solution. This gives rise to a GTSP.

6 The generalized minimum perfect matching problem

The E-GMPMP and L-GMPMP are polynomial. Indeed, the E-GMPMP remains a classical MPMP on the shrunk graph, where $c_{k\ell} := \min\{c_{ij} : i \in V_k, j \in V_\ell\}$ for $\{k, \ell\} \in E^S$. Moreover the L-GMPMP can be reduced to the E-GMPMP.

7 The generalized minimum 2-edge-connected network problem

The *Generalized Minimum Cost 2-Edge-Connected Network Problem* (E-G2ECN) consists of finding a minimum cost 2-edge-connected subgraph that contains *exactly* one vertex from each cluster (Figure 2).

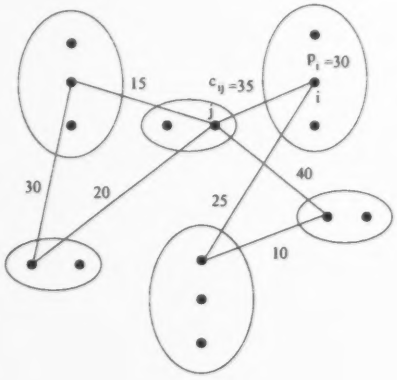

Figure 2: A feasible E-G2ECN solution

This problem arises in the context of telecommunications when copper wire is replaced with high capacity optic fiber. Because of its high capacity, this new technology allows for tree-like networks. However, this new network becomes failure-sensitive: if one edge breaks, all the network is disconnected. To avoid this situation, the network has to be reliable and must fulfill survivability conditions. Since two failures are not likely to occur simultaneously, it seems reasonable to ask for a 2-connected network.

This problem is a generalization of the GMSTP. Local networks have to be interconnected by a global network; in every local network, possible locations for a gate (location where the global network and local networks can be interconnected) of the global network are given. This global network has to be connected, survivable and of minimum cost.

The E-G2ECNP and the L-G2ECNP are studied in Huygens [29]. Even when the edge costs satisfy the triangle inequality, the E-G2ECNP and the L-G2ECNP are not equivalent. These problems are $\mathcal{NP}$-hard. There cannot exist a polynomial-time heuristic with bounded worst-case ratio for E-G2ECNP. In Huygens [29], new families of facet-defining inequalities for the polytope associated with L-G2ECNP are provided and heuristic methods are described.

8 The generalized minimum clique problem

In the *Generalized Minimum Clique Problem* (GMCP) non-negative costs are associated with vertices and edges and the graph is $|K|$-partite complete. The GMCP consists of finding a subset of vertices containing *exactly* one vertex from each cluster such that the cost of the induced subgraph (the cost of the selected vertices plus the cost of the edges in the induced subgraph) is minimized (see Figure 3).

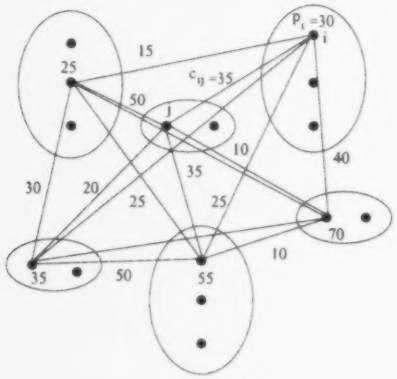

Figure 3: A feasible GMSCP solution

The GMCP appears in the formulation of particular *Frequency Assignment Problems* (FAP) (see Koster [34]). Assume that "...we have to assign a frequency to each transceiver in a mobile telephone network, a vertex corresponds to a transceiver. The domain of a vertex is the set of frequencies that can be assigned to that transceiver. An edge indicates that communication from one transceiver may interfere with communication from the other transceiver. The penalty of an

edge reflects the priority with which the interference should be avoided, whereas the penalty of a vertex can be seen as the level of preference for the frequencies."(Koster, Van Hoesel and Kolen [35]).

The GMCP can also be used to model the conformations occurring in proteins (see Althaus, Kohlbacher, Lenhof and Müller [2]). These conformations can be adequately described by a rather small set of so-called rotamers for each amino-acid. The problem of the prediction of protein complex from the structures of its single components can then be reduced to the search of the set of rotamers, one for each side chain of the protein, with minimum energy. This problem is called the *Global Minimum Energy Conformation* (GMEC). The GMEC can be formulated as follows. Each residue side chain of the protein can take a number of possible rotameric states. To each side chain is associated a cluster. The vertices of this cluster represent the possible rotameric states for this chain. The weight on the vertices is the energy associated with the chain in this rotameric state. The weight on the edges is the energy coming from the combination of rotameric states for different side chains.

The GMCP is $\mathcal{NP}$-hard (Koster, Van Hoesel and Kolen [35]). Results of polyhedral study for the GCP were embedded in a cutting plane approach by these authors to solve difficult instances of frequency assignment problems. The structure of the graph in the frequency assignment application is exploited using tree decomposition approach. This method gives good lower bounds for difficult instances. Local search algorithms to solve FAP are also investigated. Two techniques are presented in Althaus, Kohlbacher, Lenhof and Müller [2] to solve the GMEC: a "multi-greedy" heuristic and a branch-and-cut algorithm. Both methods are able to predict the correct complex structure on the instances tested.

9 The generalized Steiner tree problem

The standard generalization of the STP is the *T-Generalized Steiner Tree Problem* in its "at least" version (L-GSTP). Let $T \subseteq V$ be partitioned into clusters. The L-GSTP consists of finding a minimum cost tree of G containing *at least* one vertex from each cluster. This problem is also known as the *Group Steiner Tree Problem* or the *Class Steiner Tree Problem*. Figure 4 depicts a feasible L-GSTP solution. The L-GSTP is a generalization of the L-GMSTP since the L-GSTP defined on a family of clusters describing a partition of V is a L-GMSTP. This problem was introduced by Reich and Widmayer [52].

The L-GSTP arises in wire-routing with multi-port terminals in physical *Very Large Scale Integration* (VLSI) design. The traditional model assuming sin-

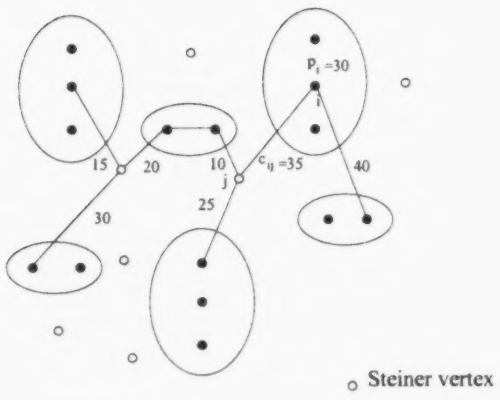

Figure 4: A feasible L-GSTP solution

gle ports for each of the terminals to be connected in a net of minimum length is a case of the classical STP. When the terminal is a collection of different possible ports, so that the net can be connected to any one of them, we have an L-GSTP: each terminal is a collection of ports and we seek a minimum length net containing at least one port from each terminal group. The multiple port locations for a single terminal may also model different choices of placing a single port by rotating or mirroring the module containing the port in the placement (see Garg, Konjevod and Ravi [21]). More detailed applications of the L-GSTP in VLSI design can be found in Reich and Widmayer [52].

The L-GSTP is $\mathcal{NP}$-hard because it is a generalization of an $\mathcal{NP}$-hard problem. When there are no Steiner vertices, the L-GSTP remains $\mathcal{NP}$-hard even if G is a tree (see Section 4). This is a major difference from the classical STP(if we assume that either there is no Steiner vertices or that G is a tree, the complexity of STP becomes polynomial). Ihler, Reich and Widmayer [31] show that the graph G can be transformed (in linear time) into a graph G' (without clusters) such that an optimal Steiner tree on G' can be transformed back into an optimal generalized Steiner tree in G. Therefore, any algorithm for the STP yields an algorithm for the L-GSTP.

Even if there exist several contributions on polyhedral aspects (see among others Goemans [24], Goemans and Myung [23], Chopra and Rao [5], [6]) and exact methods (see for instance Koch and Martin [33]) for the classical problem, only a few are known, as far as we are aware, for the L-GSTP. Polyhedral aspects are studied in Salazar [55] and a lower bounding procedure is described in Gillard and Yang [22].

A number of heuristics for the L-GSTP have been proposed. Early heuristics for the L-GSTP are developed in Ihler [30] with an approximation ratio of $|K|-1$. Two polynomial-time heuristics are tested on instances up to 250 vertices in Ihler, Reich and Widmayer [31], while a randomized algorithm with polylogarithmic approximation guarantee is provided in Garg, Konjevod, Ravi [21]. A series of polynomial-time heuristics are described in Helvig, Robins, Zelikovsky [27] with worst-case ratio of $\mathcal{O}(|K|^{\epsilon})$ for $\epsilon > 0$. These are proved to empirically outperform one of the heuristic developed in Ihler, Reich and Widmayer [31].

In the *Vertex Weighted Steiner Tree Problem* (VSTP) introduced by Segev [57], weights are associated with the vertices in V. These weights can be negative, in which case they represent profit gained by selecting the vertex. The problem consists of finding a minimum cost Steiner tree (the sum of the weights of the selected vertices plus the sum of the weights of the selected edges). This problem is a special case of the *Directed Steiner Tree Problem* (DSP) (see Segev [57]). Given a directed graph $G = (V, A)$ with arc weights, a fixed vertex and a subset $T \subseteq V$, the DSP requires the identification of a minimum weighted directed tree rooted at the fixed vertex and spanning T. The VSTP has been extensively studied (see Duin and Volgenant [13], Gorres [25], Goemans and Myung [23], Klein and Ravi [32]). As far as we know, no *Generalized Vertex Weighted Steiner Tree Problem* has been addressed. An even more general problem would be the *Vertex Weighted Directed Steiner Tree Problem.*

10 The generalized shortest path problem

Li, Tsao and Ulular [43] describe an S-generalization of the SPP in its "at most" version (M-GSPP). Let o and d be two vertices of G and assume that $V\backslash\{o, d\}$ is partitioned into clusters. The M-GSPP consists of determining a shortest path from o to d that contains *at most* one vertex from each cluster. Note that the T-generalization is of no interest since it reduces to computing the shortest paths between all the pairs of vertices belonging to the two different clusters.

In the problem considered by Li, Tsao and Ulular [43], each vertex is assigned a non-negative weight. The problem consists of finding a minimum cost path from o to d such that the total vertex weight on the path in each traversed cluster does not exceed a non-negative integer ℓ (see Figure 5). This problem with $\ell = 1$ and vertex weights equal to one for each vertex coincides with the M-GSPP.

The problem arises in optimizing the layout of private networks embedded in a larger telecommunication network. A vertex in $V\backslash\{o, d\}$ represents a digital cross connect center (DCS) that treats the information and insures the transmission. A cluster corresponds to a collection of DCS located at the same location

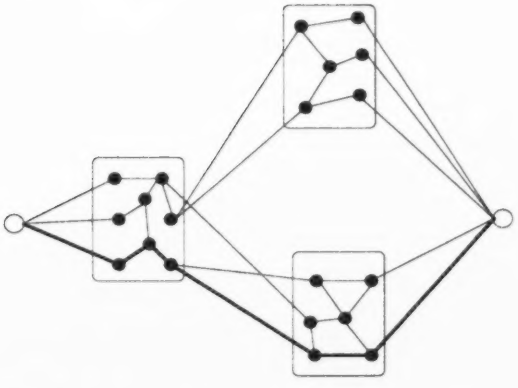

Figure 5: A feasible M-GSPP solution for $\ell = 3$

(office). The vertices o and d are two customer locations. Typically, each customer accesses the network through a single designated office. In this office, a customer can access multiple DCS. The objective in optimal network layout is to minimize the material cost. This cost includes link cost and DCS cost. Moreover, the number of DCS in each office has to be minimized in order to reduce the delay and the noise in the transmission. Li, Tsao and Ulular [43] show that this problem is polynomial in ℓ if the shrunk graph of G is series-parallel.

11 Conclusion

The generalized structures described in the previous sections can be summarized as follows. Given a graph with subsets of vertices, feasible solutions are required to satisfy an additional constraint with respect to the number of selected vertices in each subset of the family. The problem is to determine a least cost feasible solution. The generalized version is almost always harder than the classical version of the corresponding combinatorial optimization problem. The generalizations we have described are encountered in a wide range of application fields including telecommunications, transportation and biology.

Aknowledgements. This work was partially supported by the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture, by the Région Bruxelles-Capitale, by the Direction générale des Relations Extérieures de la Région wallonne de Belgique (99.13424) and by the Canadian Natural Sciences and Engineering Research Council under grant OGP0039682. The research of the second author was done in part during a visit at the Institut de Mathématiques of the Ecole Polytechnique Fédérale de Lausanne. This support is gratefully acknowledged.

References

[1] R.K. Ahuja, T.L. Magnanti, J.B. Orlin. *Network Flows: Theory, Algorithms, and Applications.* Prentice-Hall, Englewood Cliffs, NJ (1993).

[2] E. Althaus, O. Kohlbacher, H–P. Lenhof, P. Müller. *A Combinatorial Approach to Protein Docking with Flexible Side Chains.* In Proceedings of the 4th Annual International Conference on Computational Molecular Biology (RECOMB-00), R. Shamir, S. Miyano, S. Istrail, P. Pevzner, and M. Waterman (Editors), ACM Press (2000) 15–24.

[3] A. Asef-Vaziri, G. Laporte, C. Sriskandarajah. *The Block Layout Shortest Loop Problem.* IIE Transactions **32** (2000) 727–734.

[4] S. Chopra. *On the Spanning Tree Polyhedron.* Operations Research Letters **8** (1989) 25–29.

[5] S. Chopra, M.R. Rao. *The Steiner Tree Problem I: Formulations, Compositions and Extension of Facets.* Mathematical Programming **64** (1994) 209–229.

[6] S. Chopra, M.R. Rao. *The Steiner Tree Problem II: Properties and Classes of Facets.* Mathematical Programming **64** (1994) 231–246.

[7] N. Christofides. *Worst-Case Analysis of a New Heuritic for the Traveling Salesman Problem.* Management Science Research Report **388**, Carnegie Mellon University, Pittsburgh (1976).

[8] W.J. Cook, W.H. Cunningham, W.R. Pulleyblank, A. Schrijver. *Combinatorial Optimization.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York (1998).

[9] V. Dimitrijević, Z. Saric. An Efficient Transformation of the Generalized Traveling Salesman Problem into Traveling Salesman Problem on digraphs. Information Sciences **102** (1997) 105–110.

[10] M. Dror, M. Haouari. *Generalized Steiner Problems and Other Variants.* Journal of Combinatorial Optimization **4** (2000) 415–436.

[11] M. Dror, M. Haouari, J. Chaouachi. *Generalized Spanning Trees.* European Journal of Operational Research **120** (2000) 583–592.

[12] M. Dror, G. Laporte, F. V. Louveaux. *Vehicle routing with stochastic demands and restricted failures.* Zeitschrift für Operations Research **37** (1993) 273–283.

[13] C.W. Duin, A. Volgenant. *Some Generalizations of the Steiner Tree in Graphs.* Networks **17** (1987) 353–364.

[14] U. Faigle, W. Kern, P.C. Pop, G. Still, *The Generalized Minimum Spanning Tree Problem*, Working Paper, Departement of Operations Research and Mathematical Programming, University of Twente, **http://www.math.utwente.nl/dos/ormp/preprints.htm** (2000).

[15] C. Feremans. *Generalized Spanning Trees and Extensions.* Ph.D. dissertation. Université Libre de Bruxelles (2001).

[16] C. Feremans, M. Labbé, G. Laporte. *On Generalized Minimum Spanning Trees.* European Journal of Operational Research **134** (2001) 457–458

[17] C. Feremans, M. Labbé, G. Laporte. *A Comparative Analysis of Several Formulations for the Generalized Minimum Spanning Tree Problem.* Networks **39** (2002) 29–34.

[18] M. Fischetti, J.J. Salazar, P. Toth. *The Symmetric Generalized Traveling Salesman Polytope.* Networks **26** (1995) 113–123.

[19] M. Fischetti, J.J. Salazar, P. Toth. *A Branch-and-Cut Algorithm for the Symmetric Generalized Traveling Salesman Problem.* Operations Research **45** (1997) 378–394.

[20] M.R. Garey, D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, New York (1979).

[21] N. Garg, G. Konjevod, R. Ravi. *A Polylogarithmic Approximation Algorithm for the Group Steiner Tree Problem.* Journal of Algorithms **37** (2000) 66–84.

[22] P. Gillard, B. Yang. *The Class Steiner Minimal Tree Problem: A Lower Bound and Test Problem Generation.* Acta Informatica **37** (2000) 193–211.

[23] M.X. Goemans, Y.S. Myung. *A Catalog of Steiner Tree Formulations.* Networks **23** (1993) 19–28.

[24] M.X. Goemans. *The Steiner Tree Polytope and Related Polyhedra.* Mathematical Programming **63** (1994) 157–182.

[25] E. Gorres. *On the Node Weighted Steiner Tree Problem.* Ph.D. dissertation, New York University (1992).

[26] M. Grötschel, C.L. Monma, M. Stoer. *Design of Survivable Networks.* Handbooks in Operations Research and Management Science, M.O. Ball, T.L. Magnanti, C.L. Monma, G.L. Nemhauser (Editors), Elsevier, Amsterdam, Vol. 7, Chap.10. (1995) 617–672.

[27] C.S. Helvig, G. Robins, A. Zelikovsky. *An Improved Approximation Scheme for the Group Steiner Problem.* Networks **37** (2001) 8–20.

[28] A.L. Henry-Labordere. *The Record Balancing Problem : A Dynamic Programming Solution of a Generalized Traveling Salesman Problem.* RIRO B-2 (1969) 43–49.

[29] D. Huygens. *Version Généralisée du Problème de Conception de Réseau 2-Arête-Connexe.* Mémoire de Licence en Sciences Mathématiques, Faculté des Sciences, Université Libre de Bruxelles (2000).

[30] E. Ihler. *Bounds on the Quality of Approximate Solutions to the Group Steiner Tree Problem.* In Graph-Theoretic Concepts in Computer Science WG90, Lecture Notes in Computer Science **484**, Springer (1991) 109–118.

[31] E. Ihler, G. Reich, P. Widmayer. *Class Steiner Trees and VLSI-design.* Discrete Applied Mathematics **90** (1999) 173–194.

[32] P. Klein, R. Ravi. *A Nearly Best-Possible Approximation Algorithm for Node-Weighted Steiner Trees.* Journal of Algorithms **19** (1995) 104–115.

[33] T. Koch, A. Martin. *Solving the Steiner Tree Problems in Graphs to Optimality.* Networks **32** (1998) 207–232.

[34] A.M.C.A Koster. *Frequency Assignment Models and Algorithms.* Ph.D. dissertation. Proefschrift Universiteit Maastricht(1999).

[35] A.M.C.A Koster, S.P.M. Van Hoesel, A.W.J. Kolen. *The Partial Constraint Satisfaction Problem: Facets and Lifting Theorems.* Operations Research Letters **23** (1998) 89–97.

[36] M. Labbé, G. Laporte. *Maximizing User Convenience and Postal Service Efficiency in Post Box Location.* Belgian Journal of Operations Research, Statistics and Computer Science **26** (1986) 22–35.

[37] G. Laporte. *Location routing problems.* In Vehicle Routing: Methods and Studies, B.L. Golden and A.A. Assad (Editors). North-Holland, Amsterdam, (1988) 163–197.

[38] G. Laporte, A. Asef-Vaziri, C. Sriskandarajah. *Some Applications of the Generalized Traveling Salesman Problem.* Journal of the Operational Research Society **47** (1996) 1461–1467.

[39] G. Laporte, S. Chapleau, P.- É. Landry, H. Mercure. *An Algorithm for the Design of Mailbox Collection Routes in Urban Area.* Transportation Research. **23B** (1989) 271–280.

[40] G. Laporte, Y. Nobert. *Generalized Traveling Salesman Problem through n Sets of Nodes: An Integer Programming Approach.* INFOR **21** (1983) 61–75.

[41] G. Laporte, F. Semet. *Computational Evaluation of a Transformation Procedure for the Symmetric Generalized Traveling Salesman Problem.* INFOR **37** (1999) 114–120.

[42] E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys (Editors). *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization.* Wiley, New York (1985).

[43] W.–J. Li, J. Tsao, O. Ulular. *The Shortest Path with at Most l Nodes in Each of the Series/Parallel Clusters.* Networks **26** (1995) 263–271.

[44] Y.N. Lien, E. Ma, B.W.S. Wah. *Transformation of the Generalized Traveling Salesman into the Standard Traveling Salesman Problem.* Information Sciences **74** (1993) 177–189.

[45] T.L. Magnanti, L.A. Wolsey, *Optimal Trees*, Handbooks in Operations Research and Management Science, M.O. Ball, T.L. Magnanti, C.L. Monma, G.L. Nemhauser (Editors), Elsevier Science, Amsterdam, Vol. 7, Chap. 9 (1995) 503–615.

[46] A.R. Mahjoub. *Two-edge Connected Spanning Subgraphs and Polyhedra.* Mathematical Programming **64** (1994) 199–208.

[47] Y.S. Myung, C.H. Lee, D.W. Tcha. *On the Generalized Minimum Spanning Tree Problem.* Networks **26** (1995) 231–241.

[48] C.E. Noon, J.C. Bean. *A Lagrangian Based Approach for the Asymmetric Generalized Traveling Salesman Problem.* Operations Research **39** (1991) 623–632.

[49] C.E. Noon, J.C. Bean. *An Efficient Transformation of the Generalized Travelling Salesman Problem.* INFOR **31** (1993) 39–44.

[50] C.E. Noon. *The Generalized Traveling Salesman Problem.* Ph.D. dissertation. Department of Industrial and Operations Engineering. The University of Michigan, Ann Arbor (1988).

[51] P.C. Pop, W. Kern, G.J. Still. *An Approximation Algorithm for the Generalized Minimum Spanning Tree Problem with Bounded Cluster Size.* Working Paper, Departement of Operations Research and Mathematical Programming, University of Twente, **http://www.math.utwente.nl/dos/ormp/preprints.htm** (2001).

[52] G. Reich, P. Widmayer, Editors. *Beyond Steiner's Problem: A VLSI Oriented Generalization.* In Graph-Theoretic Concepts in Computer Science WG89, Lecture Notes in Computer Science **411**, Springer (1990) 196-210.

[53] J. Renaud, F.F. Boctor. *An Efficient Composite Heuristic for the Symmetric Generalized Traveling Salesman Problem.* European Journal of Operational Research **108** (1998) 571–584.

[54] J. Renaud, F.F. Boctor, G. Laporte. *A Fast Composite Heuristic for the Symmetric Traveling Salesman Problem.* INFORMS Journal on Computing **4** (1996) 134–143.

[55] J.–J. Salazar González. *A Note on the Generalized Steiner Tree Polytope.* Discrete Applied Mathematics **100** (2000) 137–144.

[56] J.P. Saskena. *Mathematical Model of Scheduling Clients through Welfare Agencies.* Journal of the Canadian Operational Research Society **8** (1970) 185–200.

[57] A. Segev. *The Node-Weighted Steiner Tree Problem.* Networks **17** (1987) 1–17.

[58] P. Slavík. *On the Approximation of the Generalized Traveling Salesman Problem.* Working Paper, Department of Mathematics, University of Buffalo (1997).

[59] L. Snyder, M.S. Daskin. *A Random-Key Genetic Algorithm for the Generalized Traveling Salesman Problem.* Working Paper, Department of Industrial Engineering and Management Sciences, Northwestern University, **http://users.iems.nwu.edu/ lsnyder/papers/gtsp.html** (2000).

[60] S.S. Srivastava, S. Kumar, R.C. Garg, P. Sen. *Generalized Traveling Salesman problem through n sets of nodes.* Journal of the Canadian Operational Research Society **7** (1969) 97–101.

[61] P. Winter. *Steiner Problem in Networks: A Survey.* Networks **17** (1987) 129–167.